SUPPLEMENTAL DATA

Biallelic CHP1 mutation causes human autosomal recessive ataxia by impairing NHE1 function

Natalia Mendoza Ferreira MSc[1], Marie Coutelier MD, PhD[2,3], Eva Janzen MSc[1], Seyyedmohsen Hosseinibarkooie PhD[1], Heiko Löhr PhD[4], Svenja Schneider MSc[1], Janine Milbradt MSc[1], Mert Karakaya MD[1], Markus Riessland PhD[1,5], Christian Pichlo MSc[6], Laura Torres-Benito PhD[1], Andrew Singleton PhD[7], Stephan Zuchner MD[8], Alexis Brice MD[2,9], Alexandra Durr MD, PhD[2,9], Matthias Hammerschmidt PhD[4], Giovanni Stevanin PhD[2,3,9], Brunhilde Wirth PhD[1]

SUPPLEMENTARY METHODS

Family and affected individuals' recruitment and evaluation

Patients and relatives from family AAR-087, together with 319 index individuals with assumed autosomal recessive cerebellar ataxia, were recruited as part of the Spastic Paraplegia and Ataxia (SPATAX) network cohort (https://spatax.wordpress.com/). Autosomal recessive inheritance was considered plausible in family AAR-087 given consanguinity and absence of clinical signs in the probands' parents. Blood samples were obtained for individuals 1, 2, 6, 7 and 8 after signed informed consent. DNA was extracted following classical procedures. Patients were examined using a standardized evaluation form (https://spatax.files.wordpress.com/2013/09/ficheclinique spataxeurospa-2011.pdf).

The girl presented with severe growth retardation (-4 SD) at age 10, exhibiting a bone age of 6 years. At age 14, primary ovarian failure was diagnosed. Immunological causes and Turner syndrome were excluded. She also had intellectual difficulties and was not able to read and write despite adapted schooling. Gait instability existed since she started walking at 12 months of age. Neurological examination at age 26 revealed spastic ataxia, with unsupported gait. Axial hypotonus and ligamentous laxity were also observed. Pyramidal syndrome was diffuse, with increased reflexes in all four limbs, bilateral extensor plantar reflex and positive Hoffman's reflex. Distal weakness was evident (2/5), mild otherwise, and associated with distal wasting. Vibration sense and pain sensation were unremarkable. She had surgical correction of *pes cavus* at age 20 and hammer toes were noticeable. Cerebellar unsteadiness and irregular finger-nose and knee-heel manoeuvres were also recognized, but no cerebellar dysarthria. Ocular motility revealed slow saccades and latencies, with limited upward and horizontal gaze. Ophthalmological examination showed normal fundus but decreased visual acuity. Hearing and smell were normal. Echocardiography was normal as well as brain MRI.

The younger brother presented with progressive gait difficulties including frequent falls at age 5. He showed moderate intellectual disability and a mild growth deficit at age 10 (-2SD).

Neurological examination at age 20 showed spastic paraparesis with cerebellar hypotonia. Increased reflexes, bilateral extensor plantar reflexes and unilaterally positive Hoffman's reflex were present. Vibration sense at ankles was decreased and distal wasting with moderate weakness in feet and hands were also registered. *Pes cavus* with hammer toes were diagnosed. Electromyography and nerve conduction studies revealed pure motor neuropathy. Ocular movements were slow, and gaze was limited upwards and horizontally with diplopia. Brain MRI showed isolated atrophy of the cerebellar vermis (Fig. 1B). Clinical features of the affected probands are summarized in (Table e-1).

Exome sequencing of family AAR-087

The Burrows-Wheeler Alignment tool (BWA[1]) and the Broad Institute's Genome Analysis Toolkit (GATK [2]), were used for alignment of sequence reads to the reference human genome (UCSC hg19) and local re-alignments around indels, base-score recalibration and variant calling, respectively. Variants were annotated and imported into GEnomes Management Application (GEM.app), a web-based tool for next generation sequencing data analysis (genomics.med.miami.edu [3]). Variants were sorted using the following criteria: (1) quality (GATK quality >50, maximal genotype quality in family >40), (2) effect on the coding sequence (missense, nonsense or splice variants, coding indels), (3) frequency in public databases <= 2% (dbSNP137) and <=1% [EVS-Exome Variant Server; NHLBI GO Exome Sequencing Project (ESP), Seattle, WA. [http://evs.gs.washington.edu/EVS/)], (4) conservation score [(GERP++ score >0 [4]], PhastCons score >0.4 [5], (5) internal frequency (less than 5 families with segregating variant amongst 892 in the exome database), (6) homozygosity in the affected patients, heterozygosity in the parents, (7) alternate allele homozygosity [Genome Aggregation Database (GenomAD [6])] and mutation effect prediction scores [SIFT [7], Polyphen-2 HDIV and HVAR [8], LRT [9] and Mutation Taster [10]] . Variants were validated, and segregation was assessed using Sanger sequencing with primers designed with Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi).

Population screening

The whole exome of 319 index cases with cerebellar ataxia was sequenced on the Illumina HiSeq 2000 as paired-end 100bp reads after library preparation with the Extended Nextera Rapid-Capture Exome kit (Illumina). Exons 2 and 6 of *CHP1* were re-sequenced due to poor coverage. Primers were designed using Primer3Plus, and amplification was performed using either the DreamTaq or the Accuprime GC rich polymerases (Thermo Scientific). Amplicons were indexed using Fast-Start High-Fidelity polymerase (Roche), then sequenced on the Illumina MiSeq using 2x300 bp paired-end read cycles. Sequence reads were processed as described above. Variants were annotated with Annovar [11] and sorted using the following criteria: (1) quality (GATK filter PASS, quality > 30), (2) effect on the coding sequence of *CHP1* (exonic non synonymous SNV or splice variant), (3) frequency in public databases <= 1% [EVS, ExAC (http://exac.broadinstitute.org/) and bioRxiv (doi: http://dx.doi.org/10.1101/030338) databases], (4) internal frequency in local exome database (allele count <= 30, ie <=4.7%).

Size exclusion chromatography

For column calibration, the following standards were used: Thyroglobulin (669 KDa), Ferritin (440 KDa), Aldolase (158 KDa), Ovalbumin (43 KDa), Carbonic anhydrase (29 KDa) and Ribonuclease A (13.7 KDa). After 48 hours of transient pcDNA3.1-CHP1-WT-V5 or CHP1-K19del-V5 transfection, cells were collected, lysed in buffer containing: 1% NP40 (AppliChem), 50 mM Tris-HCl (pH: 7.4) (Sigma-Aldrich), 150 mM NaCl (AppliChem) and 2 mM $MgCl_2$ (Sigma-Aldrich) and protease inhibitors cocktail (Roche), sonicated and centrifuged at 16000 g for 15 min. The same buffer without protease inhibitors was used for column equilibration. 250 µl of cell lysate were injected into the column and 500 µl fractions were collected. 25 µl of each fraction were processed for western blot.

Western blot, imaging and motor analysis of zebrafish

Zebrafish protein lysates for WB were prepared following the protocol described in[12], as follows: 20 fish larvae were dechorionized after, placed in soft-tissue homogenizing CK14-0.5 ml columns (PeqLab-VWR international), resuspended in 60 µl of RIPA buffer (Sigma-Aldrich) containing protease inhibitors (Roche) and homogenized in the Precellys24 (PeqLab-VWR international). Supernatants were collected after 25 min centrifugation at 16000 g and processed for WB as described above. Antibodies used are given in (table. e-4).

For CaP-MN imaging, fish were embedded in low-melting agarose micro-slides, and the first 10 motor axons posterior to the yolk sac were considered for quantification. Based on overall appearance, CaP-MN were classified as: normal, short (truncated axonal projection) or absent (total axonal atrophy). Based on terminal branching, axons were classified as normal, mild (branching ventral from midline), medium (2-3 or more branches at ventral or midline) or severe (>3 branches ventral or dorsal from midline). For cerebellar analysis, fish were dorsally embedded in channelled micro-slides. Lack of axonal density in cerebellum midline was considered as altered structure[13]. For motor analysis, ~34hpf zebrafish, still in chorion, were placed in 12-well plates containing Hank's solution and equilibrated for 5 min. Contractions in 30 sec were calculated manually for each fish. Videos were recorded with a 16MP HD (30fps) camera (Samsung) attached to a Leica S8AP0 binocular (Leica) and processed with Adobe Ae 2015.

SUPPLEMENTARY TABLESDC

Supplementary Table e-1. Sequences of primers and morpholinos used in this study.

Application	Description	Sequence
Site-directed mutagenesis	Primer	FWD: 5'-CTCGAGGAGATCAAGGAGACCGGCTTTTCC-3'
	Primer	REV: 5'-CTCGAGGAGATCAAGGAGACCGGCTTTTCC-3'
Semi-Quantitative RT-PCR	*CHP1* Primer	FWD: 5'-CTTCTTTCCAGAGGGAGAGGACCAG-'3
	CHP1 Primer	REV: 5'-CCTTCTCCAAAACCTTAACAAATTCTGTG -'3
	HPRT Primer	FWD: 5'- AAGGAGATGGGAGGCCAT -'3
	HPRT Primer	REV: 5'- GTTGAGAGATCATCT-CCACCAAT -'3
Zebrafish functional testing	*chp1* Morpholino	5'-CTGGAGCCCATGACTGCTGAAGATC-'3
	Control (non-targeting) Morpholino	5'-CCTCTTACCTCAGTTACAATTTATA-'3

Supplementary Table e-2. Antibodies used in this study

Application	Antibody	Dilution	Manufacturer
Western blot	mouse α-V5	1:5000 o/n 4°C	Thermo Scientific
	mouse α-NHE1	1:1000 o/n 4°C	Santa Cruz Biotechnology
	rabbit α-HSP90	1:1000 o/n 4°C	Cell Signalling Technology
	mouse α-Vimentin	1:1000 o/n 4°C	Sigma-Aldrich
	rabbit α-EGFR	1:1000 o/n 4°C	Santa Cruz biotechnology
	goat-α-mouse-HRP	1:2500 1h/ RT	Jackson Immuno-Research
	goat-α-rabbit-HRP	1:2500 1h/ RT	Thermo Scientific
Zebrafish western blot	rabbit α-CHP1 N-terminal (ZF)	1:1000 o/n 4°C	Aviva Systems Biology
	rabbit α-beta-actin (ZF)	1:1000 o/n 4°C	Anaspec-Eurogentec
Cells Immunostaining	rabbit α-GM130	1:500 o/n 4°C	Abcam
	mouse α-Ub	1:250 1h RT	Santa Cruz biotechnology
	mouse α-p62	1:50 1h RT	Abcam
	Phalloidine AlexaFluor 568	1:40 1h RT	Life Technologies
	mouse α-AlexaFluor-568	1:500 1h RT	Life Technologies
	rabbit α-AlexaFluor-568	1:500 1h RT	Life Technologies
	mouse α-AlexaFluor-647	1:500 1h RT	Life Technologies
	rabbit α-AlexaFluor-647	1:500 1h RT	Life Technologies
Zebrafish Immunostaining	mouse α- znp1	1:150 o/n 4°C	Hybridoma Bank
	mouse α-AlexaFluor488	1:250 o/n 4°C	Life Technologies

Supplementary table e-3. Clinical features and neurophysiological findings of the affected probands.

	♀ **Patient 6**	♂ **Patient 8**
Age at examination	26 years	20 years
Age of onset	12 months	5 years
Growth retardation	-4 SD	-2 SD
Pyramidal signs (Motor Neuron)		
Spastic paraplegia (gait difficulties)	+	+
Increased deep tendon reflexes, Babinski and Hoffmann signs	+	+
Distal muscle atrophy and weakness	+	+
Cerebellar signs		
Gait instability	+	+
Ocular dysmetria	+	+
Motor dysmetria	+	Not available
Hypotonia	Not available	+
Others		
Pes cavus, hammer toes	+	+
Sensory involvement	-	+
Mild intellectual disability	+	+
Hypergonadotropic hypogonadism	+	Not available
Brain MRI	Normal	Cerebellar vermis atrophy

Table legend: (+) and (–) symbols denote presence and absence, respectively. SD indicates standard deviation.

Supplementary table e-4. Characteristics of variants segregating in family AAR-087

Mutation Characteristics												Mutation effect prediction					Conservation	
Chr	Start	End	Ref	Alt	Gene	Aminoacid change in longest transcript	1000 Genomes Frequency	ExAC Frequency	ESP6500 Frequency	dbSNP ID	Gnom AD	Sift Score	Polyphen 2 HDIV Score	Polyphen 2 HVAR Score	LRT score	Mutation Taster score	GERP ++_RS	phyloP 100way
3	184009964	184009964	C	T	***ECE2 ****	NM_014693:c.C2590T:p.R864W	0.0002	0.0001	.	rs201775833	1	0 (D)	0.998 (D)	0.764 (P)	N	N (0.994)	2.62	0.144
3	193032845	193032848	GAG	-	***ATP13A5***	NM_198505:c.2071_2073del:p.691_691del	.	0.0002	0.0001	.	.	.	.	.	.	.	5.56	7.651
11	119045378	119045378	C	T	***NLRX1***	NM_001282144:c.C1066T:p.R356W	0.0004	0.0003	0.0003	rs142087333	0	0.07 (T)	0.999 (D)	0.96 (D)	N	D (0.994)	4.2	4.737
15	40764353	40764353	G	A	***CHST14 *****	NM_130468:c.G941A:p.R314Q	0.0002	8.24E-06	.	rs556002453	0	0.31 (T)	0.004 (B)	0.006 (B)	N	N (0.939)	1.84	0.769
15	41523634	41523637	AAG	-	***CHP1***	NM_007236:c.54_56del:p.19_19del	.	.	.	.	.	.	.	.	.	.	5.79	6.329
15	42652235	42652235	C	A	***CAPN3***	NM_000070:c.C232A:p.P78T	0.0004	0.0004	0.0004	rs138867099	0	0.01 (D)	0.614 (P)	0.524 (P)	U	D (1)	5.98	5.043
15	44127266	44127266	C	T	***WDR76***	NM_024908:c.C470T:p.S157L	.	0.0001	0.0004	rs139119504	0	0.05 (D)	0.968 (D)	0.192 (B)	N	N (0.841)	2.73	2.702
15	45412435	45412435	G	A	***DUOXA1 ****	NM_144565:c.C638T:p.T213M	0.0036	0.0039	0.0024	rs149960164	11	0 (D)	1 (D)	0.995 (D)	D	D (1)	5.12	6.445
15	57999082	57999082	G	C	***POLR2M***	NM_015532:c.G42C:p.E14D	0.0006	0.0015	0.0015	rs150413697	0	0.39 (T)	0.998 (D)	0.941 (D)	.	N (1)	3.84	3.204
15	83933230	83933230	C	T	**BNC1**	NM_001717:c.G773A:p.G258E	0.004	0.0008	0.0023	rs116464429	2	0.01 (D)	1 (D)	0.986 (D)	D	D (1)	5.75	5.842
15	86312927	86312927	G	A	***KLHL25 *****	NM_022480:c.C115T:p.L39F	.	.	.	.	.	0.21 (T)	0.669 (P)	0.512 (P)	N	D (0.992)	3.93	3.297
18	2769780	2769780	C	T	***SMCHD1 *****	NM_015295:c.C4808T:p.T1603I	0.004	0.0012	0.0026	rs147034750	2	0.2 (T)	0 (B)	0 (B)	N	N (1)	0.292	-0.226

Table legend: Variant annotation was performed using Annovar (www.openbioinformatics.org/annovar/). Variant frequencies were looked for in 1000 Genomes database (http://www.1000genomes.org/), Exome Aggregation Consortium (ExAC, http://exac.broadinstitute.org/), Exome Variant Server (EVS, http://evs.gs.washington.edu/EVS/), and identifiers were researched in dbSNP (http://www.ncbi.nlm.nih.gov/SNP/). Pathogenicity prediction scores are estimated as follows: SIFT (http://sift.jcvi.org/) predicts deleteriousness (D) under 0.05, otherwise the variant is assumed to be tolerated (T); Polyphen2 (http://genetics.bwh.harvard.edu/pph2/) classifies SNPs as probably damaging (D; HDIV>=0.957, HVAR>=0.909), possibly damaging (P; 0.453<=HDIV<=0.956, 0.447<=pp2_hdiv<=0.908), or benign (B; HDIV<=0.452, HVAR <=0.446); LRT differentiates variants with deleterious (D), neutral (N) or unknown (U) effect; MutationTaster (http://www.mutationtaster.org/) classifies them as "disease_causing_automatic" (A), "disease_causing" (D), "polymorphism" (N) or "polymorphism_automatic" (P) with a given probability value, 1 being the most probable. For both GERP++ (http://mendel.stanford.edu/SidowLab/downloads/gerp/) and PhyloP (http://compgen.bscb.cornell.edu/phast/help-pages/phyloP.txt), higher scores indicate better residue conservation. * indicates: excluded for homozygosity in alternate allele (GnomAD). ** indicates: Excluded for weak predicted effect (4/5 or 5/5 scores)

Supplementary table e-5. Focused screening of *CHP1* mutations in ARCA cohort.

Gene	Exon rank in transcript NM_007236	Length	Target (hg19)	Average coverage by WES	Average percentage above 10 by WES	Average percentage above 30 by WES	Average coverage by panel	Average percentage above 10 by panel	Average percentage above 30 by panel
CHP1	1	66	chr15:41523582-41523647	58.8	98.2	89.8			
CHP1	2	73	chr15:41535865-41535937	14.6	73.8	6.0	4943	99.5	99.1
CHP1	3	81	chr15:41549108-41549188	74.5	98.0	89.3			
CHP1	4	128	chr15:41554954-41555081	46.0	98.9	75.3			
CHP1	5	62	chr15:41562755-41562816	57.3	99.6	93.3			
CHP1	6	123	chr15:41570965-41571087	25.2	94.2	25.4	305	98.1	94.6
CHP1	7	54	chr15:41571534-41571587	42.2	99.2	81.3			
	All exons			**44.0**	**94.7**	**62.2**			
								Total with both methods	
								98.7	**88.3**

Table legend: Coverage and percentage of bases covered at 10x and 30x for all *CHP1* exons by WES. WES average coverage of the ARCA cohort is 44x. Excellent coverage levels were reached upon exons 2 and 6 resequencing, with 100% 30x coverage in 99.1% (n=316) and 91.8% (n=293) of the patients respectively, and a mean coverage of 4943x and 305x, respectively.

SUPPLEMENTARY VIDEOS

Supplementary video 1. Spontaneous contractions of uninjected zebrafish (left panel) and *chp1* morphants (right panel) analyzed at ~32hpf.

Supplementary video 2. Abnormal fin coiling without locomotion of *chp1* morphant analyzed at ~72hpf

SUPPLEMENTARY REFERENCES

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-1760.
2. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 2010;20:1297-1303.
3. Gonzalez MA, Lebrigio RF, Van Booven D, et al. GEnomes Management Application (GEM.app): a new software tool for large-scale collaborative genome analysis. Hum Mutat 2013;34:842-846.
4. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS computational biology 2010;6:e1001025.
5. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research 2005;15:1034-1050.
6. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-291.
7. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research 2003;31:3812-3814.
8. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics 2013;Chapter 7:Unit7 20.
9. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome research 2009;19:1553-1561.
10. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nature methods 2010;7:575-576.
11. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 2010;38:e164.
12. Hoyt C. Rapid, high-throughput homogenization of embryonic or larval zebrafish (*Danio rerio*).Protocol Exchange 2009. http://dx.doi.org/10.1038/nprot.2009.211.
13. Borck G, Hog F, Dentici ML, et al. BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. Genome research 2015;25:155-166.